



Wilson, R. E. (2008). *From inductance loops to vehicle trajectories*.
<http://hdl.handle.net/1983/1162>

Early version, also known as pre-print

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

FROM INDUCTANCE LOOPS TO VEHICLE TRAJECTORIES

R. Eddie Wilson

Department of Engineering Mathematics, University of Bristol, Queen's Building, University Walk, BRISTOL BS8 1TR, United Kingdom

Tel.: +44-117-928-8967

Fax.: +44-117-954-6833

Email: RE.Wilson@bristol.ac.uk

Submitted 1st August 2008 to the TRB annual meeting.

4118 words text, 2 tables, 4 figures (5668 words equivalent < 7500 limit)

Abstract

This paper describes ongoing work on the United Kingdom's M42 motorway which has a uniquely high coverage of inductance loop detectors. The spacing of detectors is sufficiently small for one to use individual vehicle data to follow single vehicles down the highway. The paper gives a brief outline of the data collection work and sketches how the vehicle re-identification algorithms work. Sample data sets are available from the project web-site <http://www.enm.bris.ac.uk/trafficdata>.

FROM INDUCTANCE LOOPS TO VEHICLE TRAJECTORIES

by R. Eddie Wilson, University of Bristol, United Kingdom

Abstract. This paper describes ongoing work on the United Kingdom's M42 motorway which has a uniquely high coverage of inductance loop detectors. The spacing of detectors is sufficiently small for one to use individual vehicle data to follow single vehicles down the highway. The paper gives a brief outline of the data collection work and sketches how the vehicle re-identification algorithms work. Sample data sets are available from the project web-site <http://www.enm.bris.ac.uk/trafficdata>.

1. BACKGROUND

Our aim is the collection of vehicle trajectory data from highways for the better understanding of traffic flow dynamics and the better calibration of micro-simulation models. One path to this goal (1,2) is to mount video cameras on tall buildings or gantries and apply computer vision techniques to the resulting video streams. However, a limitation is that the resulting data sets are quite small (order 15 minutes is typical) — because a great deal of manual intervention is required for the computer vision algorithms to work correctly.

An alternative source of traffic data is the inductance loop detection infrastructure which is a common feature of highways in the developed world. The spacing between loop detectors varies, but in Western Europe it is typically 500m (approx 1650 feet) and in the United States it is similar. Loop detectors may either be single, in which case they measure only flow and occupancy, or double, in which case they also measure vehicles' velocities and lengths. In their usual operation, these measurements are bundled into time averages (1 minute is a typical unit) and communicated back to a control office. The consequent spatiotemporal data has led to an intense discussion of macroscopic traffic patterns and the fundamental mechanisms which explain them (3,4,5,6).

The focus of our ongoing work is Individual Vehicle Data (IVD) collected from inductance loop detectors. For this, one intercepts and stores the velocity, length and timing information of individual vehicles before the time-average is applied. Some past studies of IVD have focused on data from one detector and the rich connections between headway statistics and lattice gases (see (7) and many subsequent papers). In contrast, Coifman and collaborators have pioneered vehicle re-identification techniques which apply pattern-matching methods to the IVD from a pair of detectors (see (8); (9) for the latest perspective). The chief idea is that a vehicle's velocity at the upstream detector may be used to forecast its arrival time at the downstream detector. One then searches for a downstream length record which matches that recorded upstream and for which the arrival time is consistent. Unfortunately, with the typical loop spacing of order 500m, traffic may shuffle significantly between detectors. Since length measurements are noisy, one in practice may only re-identify a vehicle with confidence if its length is sufficiently distinguished or if the traffic flow is sufficiently light. Coifman has thus been limited by the spatial resolution of his data and has worked mostly on the re-identification of trucks for the purpose of monitoring segment journey times.

The new opportunity described here is provided by the English Highways Agency's (10) Active Traffic Management (ATM) system which operates on a 15km (approx 9 mile) stretch of the M42 motorway — constituting part of the 'box' of motorways around Birmingham (the United

Kingdom's second largest city). In busy periods, variable message signs set reduced speed limits and open the emergency breakdown lane for ordinary driving. Because of the need to monitor traffic closely in this situation, inductance detectors have been installed much more densely than is usual — with a nominal spacing of 100m (approx 330 feet). However, in a 900m section where queuing is common, this spacing is reduced to circa 30m (approx 100 feet). In normal operation, the ATM system captures the usual 1-minute average data, but the spatial resolution is such that we may examine the structure of stop-and-go waves in a level of detail that was not previously possible, see fig. 1.

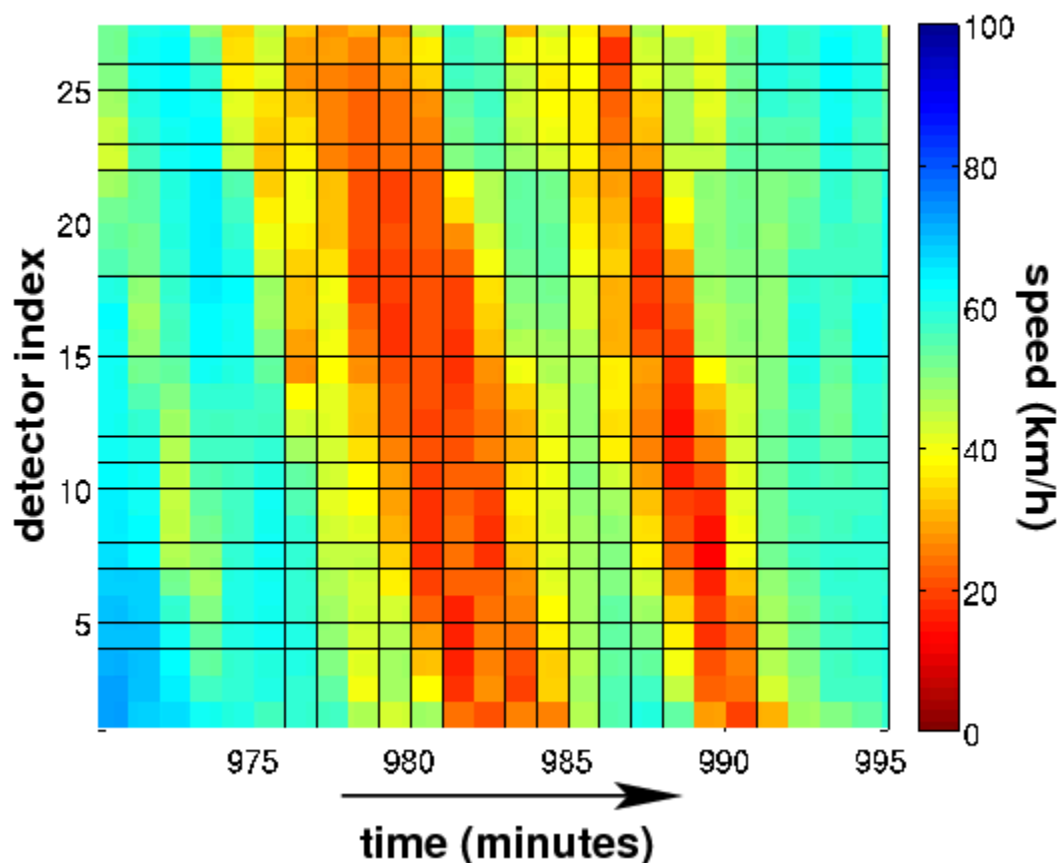


FIGURE 1 A spatio-temporal plot of speed (averaged across 3 running lanes) for the high coverage section of M42 ATM showing two stop-and-go waves. No interpolation has been used in the production of this picture. The vertical extent is approximately 900m (about 3000 feet). We believe that this degree of spatial resolution is unique.

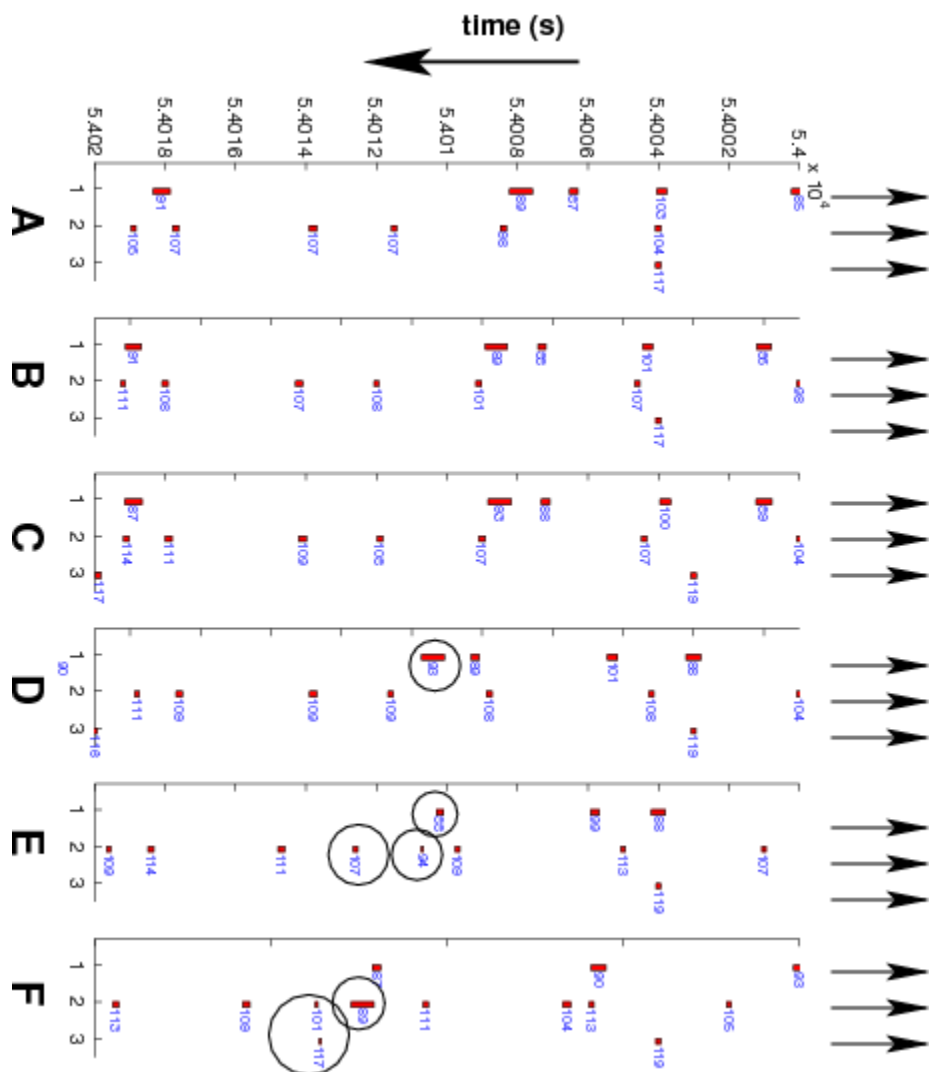


FIGURE 2 Visualisation of 20 seconds of IVD captured during the 2003 trial. Panels A–F denote the six detector sites progressing in downstream order. In each panel, lane numbers 1–3 are plotted horizontally whereas time is plotted down the vertical axis and thus plays the role of a space-like coordinate in which vehicles drive up the axis (rather like a photo-finish camera). Vehicle records are illustrated by rectangles (whose size is derived from the vehicle’s length) next to which the velocity in km/h is given. There is a time-offset of 3 seconds between each panel so that vehicles at 120km/h (approx 75 miles per hour) maintain the same horizontal level. Vehicles may clearly be re-identified from panel to panel and the overall effect is similar to 6 ‘helicopter-views’ of the traffic showing how the relative configuration of vehicles changes down the highway. Note that in the UK, slow traffic (trucks etc.) drives on the left. Some lane-changing events, where vehicles straddle detectors, have been circled.

In 2003, transport engineering consultancy TRL conducted a 2-day trial where IVD was collected from 6 consecutive ATM inductance detectors with a nominal spacing of 100m. When displayed through an appropriate graphical user interface (see fig. 2), the intelligent human can identify the patterns of vehicles from detector to detector and thus it appears possible to re-identify almost all vehicles with confidence — not just between a pair of detectors, but through the entire 500m section. Thus in effect one coarsely re-constructs the trajectories of vehicles (although fine details of vehicles' accelerations cannot be captured). The technical challenge is then to devise algorithms which replicate the human pattern-matching process (11). In this respect, we developed algorithms which re-identify vehicles over 100m with a success rate which on average exceeds 99% (using a human-matched set as the ground truth). Unfortunately, this success rate is due in part to the anomalously quiet traffic conditions experienced during the trial (in particular flow-breakdown did not occur). Since ongoing improvements in communications hardware and standards have reduced the need to access road-side hardware for IVD capture, there is now scope for a much more comprehensive data collection exercise.

In the work that we announce here, we have exploited a commercial equipment trial to capture the IVD from 16 consecutive detectors over 1.5km (nearly one mile) of the North-bound M42. The trial runs from January to October 2008, and since week-day traffic flows are approximately 70,000 vehicles, each detector will capture and store the IVD of over 15,000,000 vehicles in total. This data resource is being made available to the traffic research community over the summer and autumn of 2008, see (12). The aims of this paper are two-fold: 1. To give further details of the data collection exercise and the basic calibration work (section 2) and 2. to give a brief outline of how the re-identification algorithms are being developed (section 3). Finally section 4 presents conclusions and lists some ideas for possible joint projects.

2. DETAILS OF THE INSTRUMENTED SECTION

We now give details of the ongoing IVD collection exercise. A kml file is available for download (13) which may be imported into Googlemaps in order to display the instrumented section, which is approximately 1.5km (nearly one mile) long. Fig. 3 gives a schematic diagram of the layout. Note that the sections of lane 0 labeled ATM are emergency breakdown lane in which vehicles do not normally drive. The Active Traffic Management system may open this lane for ordinary driving in peak periods, but at present this facility is not used within the instrumented section.

The key feature in fig. 3 is the mid-section on-ramp, which contributes order 10–15% of the downstream traffic on average, divided roughly equally between its two lanes. Firstly, this presents new challenges for the vehicle re-identification algorithm between detectors 4 and 9 where the majority of merges occur. Secondly, it introduces the possibility of interesting traffic dynamics. Indeed, an examination of 1-minute average data (which has been collected for several years) indicates that this section regularly induces flow break-down and stop-and-go waves as well as experiencing large amplitude stop-and-go waves which have propagated back from highly congested junctions further downstream.

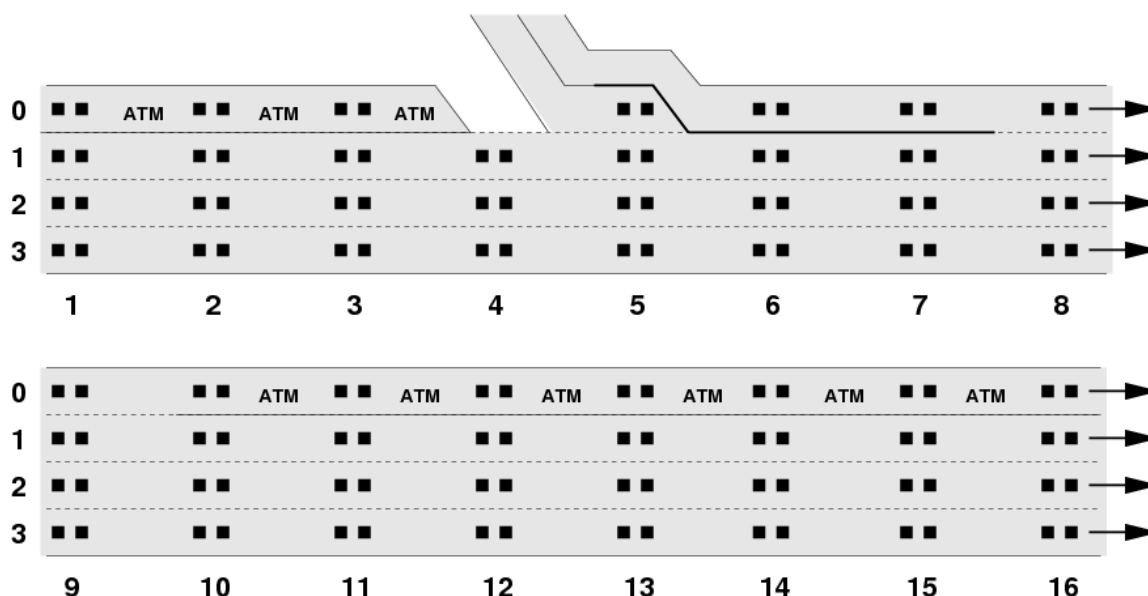


FIGURE 3 Schematic layout of the instrumented section of motorway. Loop detector sites 1–16 are separated by about 100m (approx 330 feet) longitudinally so that the whole section is approximately 1.5km (nearly one mile). Lanes 1–3 constitute ordinary running lanes to which UK driving rules apply (slower vehicles should tend to drive on the left in lane 1). Lane 0 is made up of the emergency breakdown lane (labeled ATM, since it may be activated as an ordinary running lane) or parts of the on-ramp. The bold lane-marking on the on-ramp denotes a region of chevrons which vehicles should not normally cross.

The data that each inductance detector records is listed in Table 1. A significant advance since our 2003 exercise is that arrivals are now quoted to 0.1 seconds accuracy (whereas previously they were quoted only to the nearest second, with in-lane headway quoted to 0.1 seconds). These new data are extracted as a by-product of the IDRIS waveform analysis system which is undergoing a commercial trial. It is a general observation that IDRIS appears to produce significantly cleaner IVD than the standard hardware used in the 2003 exercise. In particular, vehicles miscounts (due to either close-following or lane-straddling) are fewer than 1 in 1000.

TABLE 1 Data fields and nominal resolution for each individual vehicle record. Note that the velocity and length records are less accurate than their nominal resolution.

Quantity	arrival time	lane	velocity	length
Resolution	0.1 s	integer 0-3	0.01 m/s	0.01 m

Our first challenge is to calibrate the loop data. Whereas we may assume that vehicles' arrival times are correct to their quoted resolution, the velocity and length records are subject to

measurement error. Furthermore, the spacing of consecutive detectors is only nominally 100m and must be determined more accurately for the re-identification algorithms to work efficiently. Note that the question of velocity calibration can only be solved definitively by driving a probe vehicle repeatedly through the instrumented section, and cross-checking with its IVD — which we have yet to do.

In contrast, the calibration of length measurements, at least in relative terms, may be achieved directly from IVD. The technique is to consider consecutive pairs of detectors and to seek upstream-downstream pairs of vehicle records which must match, because conditions are sufficiently quiet that there are no other vehicle arrivals at either detector within any reasonable time tolerance. This requirement can be specified precisely and many thousands of such unique possible matches can be found during the night. Joint distributions of the pairs of measured lengths can then be analyzed. For private cars (say length less than 5.5m, approx 18 feet) the difference in length measurements between consecutive detectors is small — with mean order 1cm (less than half an inch) and standard deviation order 7cm (approx 3 inches). Consequently, detectors' length measurements do not require calibration and moreover there is sufficient information in them to assist the re-identification of even private cars (whose lengths are not especially distinguished). Measured length-differences for trucks are more widely spread but this is not a serious problem because of their relative scarcity. Unfortunately trucks' statistics do contain many outliers due to lane changing-events, where an anomalous length is recorded due to the straddling of detectors in adjacent lanes.

Unique possible matches can also be used to discover the true driving distance between detectors, by taking the time-difference between upstream and downstream records and multiplying by the mean of their velocities. This calculation gives a distribution of distances with a spread which is due principally to the ± 0.1 s accuracy of the time-difference. By taking a large number of records, the true driving distance may be extracted from the statistics, see Table 2. Since this method assumes accurate velocity measurements, we checked the Table 2 distances with Googlemaps and found very close agreement. This indicates that the errors in velocity measurement have a very small mean component.

TABLE 2 Actual driving distances between consecutive pairs of detector sites.

loops	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9
gap (m)	106.0	99.5	101.6	91.4	98.8	92.4	101.7	91.3

loops	9-10	10-11	11-12	12-13	13-14	14-15	15-16
gap (m)	92.7	92.9	79.1	109.9	103.2	98.6	99.9

Now that driving distances have been determined accurately, we may use them to seek unique possible matches with much tighter tolerances than we used previously. We thus dramatically increase the number of unique possible matches which in turn leads to more accurate calibration of the loops. The drawback of this boot-strapping approach is that length-error statistics are built on

portions of traffic data where there are lots of unique possible matches — that is, principally sparse and hence fast moving traffic. Thus there are potential limitations in exploiting the length-error statistics to re-identify slow moving traffic.

3. OUTLINE OF RE-IDENTIFICATION ALGORITHMS

We now sketch how the re-identification algorithms work. (The full details of the algorithms will be the subject of a forthcoming journal paper.) For simplicity, we may focus on matching the records for a single pair of detectors which do not overlap with the on-ramp, for example, numbers 11 and 12. (The matching for detector pairs 12–13, 13–14, 14–15 and 15–16 is rather similar.) In fact, algorithms under development make use of the information from more than two detectors simultaneously, but are beyond the scope of this discussion. Since the re-identification of sparse traffic turns out to be rather trivial, it is necessary to test algorithms on a sufficiently congested day with strongly dynamic traffic patterns, for example 24th January 2008. See fig. 4.

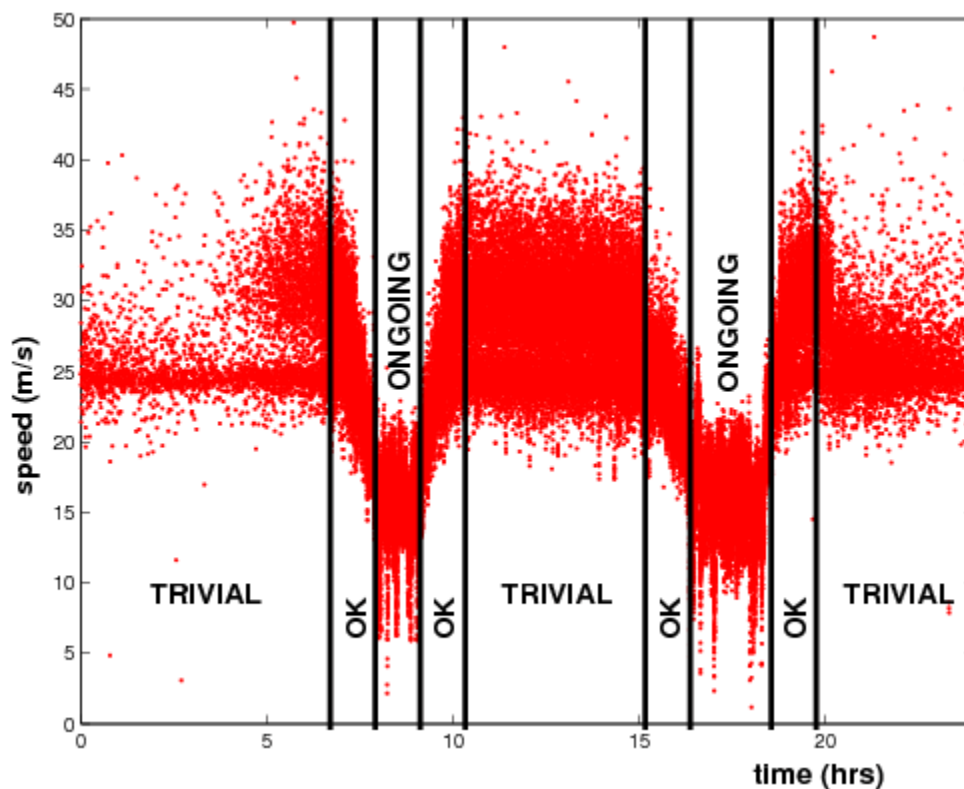


FIGURE 4 Individual vehicle velocity data from detector 11 for 24th January 2008. In summary, we are presently able to re-identify traffic robustly except in the periods 08:00–09:30 hrs and 16:30–18:30 hrs when the mean speed drops sharply and stop-and-go waves nucleate at this location. Our success includes the period 07:00–08:00 hrs when the flow rate at times approaches 6000 vehicles/hour summed over the three running lanes. During 07:00–08:00 hrs we observe a small velocity variance measured both temporally and between lanes.

The first step in re-identification is to partition the data so that when matches are sought, only relatively small numbers of vehicles need to be analyzed simultaneously. This partition exploits the possible match idea which we introduced in section 2. Specifically, for each vehicle record at the upstream detector, we forecast the arrival time at the downstream detector using the distance x between the detectors (determined precisely in section 2) and the velocity and arrival time at the upstream detector. We then find all records at the downstream detector whose actual arrival time is within a tolerance e of the forecast. The tolerance e may be designed in various ways, but must include 0.1s to account for the nominal error in time measurements, as well as other components which model for velocity measurement error or the possibility of non-zero acceleration.

We then apply the procedure in reverse: that is, for each downstream record we forecast the earlier upstream arrival time based on downstream velocity and find all possible matches at the upstream detector. In this way we construct a bipartite graph of connections between upstream and downstream vehicle records. The data is then partitioned into sets of possible matches by finding the maximal connected components of the (symmetrized) bipartite graph. We call these match-sets. For example, a unique possible match corresponds to a match-set with a single upstream and a single downstream record. In busy traffic, match-sets grow in size, because headways are smaller, so that one finds more and more vehicles within any given tolerance interval. Note that large match-sets tend to make re-identification more complicated and more computationally intensive, but they do not necessarily render it intractable.

We now consider the analysis of a single match-set. The simplest situation is the ‘square’ case where it consists of an equal number of upstream and downstream records. To re-identify we then seek the ‘best’ bijection between the upstream and downstream records. The simplest version of this technique defines a pairwise error score with (i) a component based on the compatibility of the upstream and downstream arrival times and velocities with the spacing between the detectors and (ii) a component based on the difference of measured vehicle lengths upstream and downstream (whose design may be informed by the joint probably distribution of measured lengths for unique possible matches). The best bijection can then be defined as that which minimizes some norm of the score vector — in the case of the 1-norm, this problem may be solved by a standard numerical technique known as the Hungarian algorithm.

In practice however, pairwise scoring methods are confused by groups of vehicles with similar velocities and lengths who pass a detector at about the same time but in different lanes. Hence more sophisticated algorithms use lane information and work on the (guessed) relative likelihood of different re-orderings of the vehicles. We developed these methods commercially during our original study (11) and tuned their parameters to perform optimally against human-matched sets. (Recall the clear patterns in fig. 2.) When compared to further human-matched data, these methods exceeded a 99% correct re-identification rate. Early indications are that this success rate is equaled in the 2008 data in all but the busiest and most strongly dynamic periods.

We now consider how the algorithms break in busy conditions. In addition to an explosion in the size of match-sets, a common problem we encounter is that match-sets are not square. Occasionally this is because a vehicle has straddled detectors during a lane-change and hence has been counted twice at one site or missed completely at the other. This type of problem is in fact relatively mild and we have a variety of ad hoc solutions for dealing with it, although the merge section still presents challenges.

More seriously, in strongly dynamic traffic conditions, we presently obtain many non-square

match-sets — this indicates that the search for possible matches is itself breaking down. This is because under harsh braking conditions, the velocity at the upstream detector does not give a reliable forecast of when the vehicle will arrive downstream. To solve this problem we are developing alternative methods which rely principally on matching the sequences of vehicle lengths recorded at each detector. However, in very slow traffic this technique will be subject to uncertainty for two reasons. Firstly, the vehicles have time to substantially re-order themselves over 100m if they are driving slowly enough. Secondly, the detectors themselves do not capture lengths reliably at very low speeds. In our favor, the behavior of slow queuing traffic is not of especial interest.

4. CONCLUSIONS

We have summarized ongoing work in the collection and re-identification of Individual Vehicle Data from inductance loops. This project uses the Active Traffic Management section of the M42 motorway, which is one of the most densely instrumented highways in the world. The section from which we collect data includes a merge and is a good location for observing complex spatiotemporal patterns in detail.

At present, our re-identification algorithms work extremely well (>99% accurate) in all but the most congested and most strongly dynamic conditions. Our current work is focused on extending the algorithms to deal with these challenging situations and moreover to identify vehicle merges correctly.

By the end of our project in October 2008, we will have constructed the trajectories of in excess of 15,000,000 vehicles. This means that for each vehicle, we will determine a 16x4 array detailing its arrival time, speed, lane number and measured length at each of the 16 detector sites, and supply pointers to the arrays of its immediate neighbors. Small samples of the data will be freely available at <http://www.enm.bris.ac.uk/trafficdata> and the complete data set will be provided on application.

Our data unfortunately cannot describe dynamics which occur between the detectors (nominal spacing 100m, approx 330 feet) and so in particular, the fine details of vehicles' accelerations are not directly accessible. However the advantage compared to camera trajectory data is the sheer volume of our data set. Consequently, statistical inference might be used to develop descriptions of driver behavior which are much more detailed than the data appears to allow at first sight. In this respect it will be interesting to see to what extent camera trajectory information and inductance loop IVD can be fused.

With such a large volume of data, we may disaggregate in many different ways and yet retain statistically significant numbers of vehicle trajectories. Thus it seems that (highly-parametrized) models of lane-changing would benefit in particular, but one could also model the dependence of driver behavior on more exotic factors such as the weather. We are open to suggestions for joint projects that take the applications of this work forward.

ACKNOWLEDGMENT

This research is supported by an EPSRC Advanced Research Fellowship (grant no. EP/E055567/1). Thanks also to Diamond Consulting, TRL and the English Highways Agency for

access to inductance loop data collected from the M42 ATM system.

REFERENCES

1. NGSIM (Next Generation Simulation) project. <http://www.ngsim.fhwa.dot.gov>
2. Hoogendoorn, S. P., H. J. Zuylen, B. G. H. van Gorte and M. G. Vosselman. Microscopic traffic data collection by remote sensing. *Transportation Research Record*, Vol. 1855, 2003, pp. 121–128.
3. Daganzo, C. F., M. J. Cassidy and R. L. Bertini. Possible explanations of phase transitions in highway traffic. *Trans. Res. A*, Vol. 33, 1999, pp. 365–379.
4. Kerner, B. S. *The Physics of Traffic*. Springer-Verlag, 2004.
5. Schönhof, M. and D. Helbing. Empirical features of congested traffic states and their implications for traffic modeling. *Transport. Sci.*, Vol. 41, 2007, pp. 135–166.
6. Wilson, R. E. Mechanisms for spatio-temporal pattern formation in highway traffic. *Phil. Trans. R. Soc. A*, Vol. 366, 2008, pp. 2017–2032.
7. Neubert, L., L. Santen, A. Schadschneider and M. Schreckenberg. Single-vehicle data of highway traffic: A statistical analysis. *Phys. Rev. E*, Vol. 60, 1999, pp. 6480–6490.
8. Coifman, B. and M. Cassidy. Vehicle reidentification and travel time measurement on congested freeways. *Trans. Res. A*, Vol. 36, 2002, pp. 899–917.
9. Coifman, B. and S. Krishnamurthy. Vehicle reidentification and travel time measurement across freeway junctions using the existing detector infrastructure. *Trans. Res. C*, Vol. 15, 2007, pp. 135–153.
10. Highways Agency. <http://www.highways.gov.uk>
11. Lunt, G., M. Day and R. E. Wilson. Enhancing motorway traffic data with novel vehicle reidentification algorithms. Proceedings of the 13th World Congress on Intelligent Transport Systems and Services (2006) 11 pages.
12. M42 Individual Vehicle Data project web-site. <http://www.enm.bris.ac.uk/trafficdata>
13. Location of inductance loops involved in the IVD trial. KML file downloadable from http://www.enm.bris.ac.uk/staff/rew/kml/m42_ivd_loops.kml